

## Claims

### WHAT IS CLAIMED IS:

#### 1. A method comprising:

selecting a modeling parameter from a plurality of modeling parameters  
characterizing a mixture of Student distribution components;

computing a tractable approximation of a posterior distribution for the  
selected modeling parameter based on an input set of data and a current estimate  
of a posterior distribution of at least one unselected modeling parameter in the  
plurality of modeling parameters;

computing a lower bound of a log marginal likelihood as a function of  
current estimates of the posterior distributions of the modeling parameters, the  
current estimates of the posterior distributions of the modeling parameters  
including the computed tractable approximation of the posterior distribution of the  
selected modeling parameter; and

generating a probability density modeling the input set of data, the  
probability density including the mixture of Student distribution components, the  
mixture of Student distribution components being characterized by the current  
estimates of the posterior distributions of the modeling parameters, if the lower  
bound is satisfactorily optimized.

2. The method of claim 1 wherein the computing operations comprise a  
first iteration and further comprising:

selecting a different modeling parameter from the plurality of modeling  
parameters and repeating in a subsequent iteration the operations of computing a

1 tractable approximation and computing a lower bound using the newly selected  
2 modeling parameter, if the lower bound is not satisfactorily optimized in the first  
3 iteration.

4 3. The method of claim 1 wherein computing a lower bound comprises:  
5 computing the lower bound of the log marginal likelihood as a function of  
6 prior distributions of the modeling parameters.

7  
8 4. The method of claim 1 wherein computing a tractable approximation of  
9 a posterior distribution comprises:  
10 computing a variational approximation of the posterior distribution of the  
11 selected modeling parameter.

12 5. The method of claim 1 wherein one of the plurality of modeling  
13 parameters represents a mean of each of the Student distribution components.

14  
15 6. The method of claim 1 wherein one of the plurality of modeling  
16 parameters represents a precision matrix of the Student distribution components.

17 7. The method of claim 1 wherein one of the plurality of modeling  
18 parameters represents a labeling parameter of the Student distribution components.

19  
20 8. The method of claim 1 wherein one of the plurality of modeling  
21 parameters represents a scaling parameter of a precision matrix of the Student  
22 distribution components.

1           9. The method of claim 1 wherein one of the plurality of modeling  
2 parameters represents a mixing coefficients parameter of the Student distribution  
3 components.

4           10. The method of claim 1 wherein generating a probability density  
5 comprises:

6           generating the probability density including the mixture of Student  
7 distribution components, the mixture of Student distribution components being  
8 characterized by the current estimates of the posterior distributions of the  
9 modeling parameters and an estimate of the number of degrees of freedom of each  
10 Student distribution component.

11           11. The method of claim 1 further comprising:

12           storing the current estimates of the posterior distributions of the modeling  
13 parameters in a storage location.  
14

15           12. The method of claim 1 wherein the input set of data represents auditory  
16 speech data from an unknown number of speakers, and further comprising  
17 determining a correct number of speakers from the probability density modeling  
18 the input set of data.

19           13. The method of claim 1 wherein the input set of data represents image  
20 segmentation data from images having regions of different characteristics.  
21  
22  
23  
24  
25

1 14. A computer program product encoding a computer program for  
2 executing on a computer system a computer process, the computer process  
3 comprising:

4 selecting a modeling parameter from a plurality of modeling parameters  
5 characterizing a mixture of Student distribution components;

6 computing a tractable approximation of a posterior distribution for the  
7 selected modeling parameter based on an input set of data and a current estimate  
8 of a posterior distribution of at least one unselected modeling parameter in the  
9 plurality of modeling parameters;

10 computing a lower bound of a log marginal likelihood as a function of  
11 current estimates of the posterior distributions of the modeling parameters, the  
12 current estimates of the posterior distributions of the modeling parameters  
13 including the computed tractable approximation of the posterior distribution of the  
14 selected modeling parameter; and

15 generating a probability density modeling the input set of data, the  
16 probability density including the mixture of Student distribution components , the  
17 mixture of Student distribution components being characterized by the current  
18 estimates of the posterior distributions of the modeling parameters, if the lower  
19 bound is satisfactorily optimized.

20  
21 15. The computer program product of claim 14 wherein the computing  
22 operations comprise a first iteration and further comprising:

23 selecting a different modeling parameter from the plurality of modeling  
24 parameters and repeating in a subsequent iteration the operations of computing a  
25 tractable approximation and computing a lower bound using the newly selected

1 modeling parameter, if the lower bound is not satisfactorily optimized in the first  
2 iteration.

3 16. The computer program product of claim 14 wherein computing a lower  
4 bound comprises:

5 computing the lower bound of the log marginal likelihood as a function of  
6 prior distributions of the modeling parameters.

7  
8 17. The computer program product of claim 14 wherein computing a  
9 tractable approximation of a posterior distribution comprises:

10 computing a variational approximation of the posterior distribution of the  
11 selected modeling parameter.

12 18. The computer program product of claim 14 wherein one of the plurality  
13 of modeling parameters represents a mean of each of the Student distribution  
14 components.

15  
16 19. The computer program product of claim 14 wherein one of the plurality  
17 of modeling parameters represents a precision matrix of the Student distribution  
18 components.

19 20. The computer program product of claim 14 wherein one of the plurality  
20 of modeling parameters represents a labeling parameter of the Student distribution  
21 components.

22  
23 21. The computer program product of claim 14 wherein one of the plurality  
24 of modeling parameters represents a scaling parameter of a precision matrix of the  
25 Student distribution components.

1           22. The computer program product of claim 14 wherein one of the plurality  
2 of modeling parameters represents a mixing coefficients parameter of the Student  
3 distribution components.

4           23. The computer program product of claim 14 wherein generating a  
5 probability density comprises:

6           generating the probability density including the mixture of Student  
7 distribution components, the mixture of Student distribution components being  
8 characterized by the current estimates of the posterior distributions of the  
9 modeling parameters and an estimate of the degrees of freedom of each Student  
10 distribution component.

11           24. The computer program product of claim 14 wherein the computer  
12 process further comprises:

13           storing the current estimates of the posterior distributions of the modeling  
14 parameters in a storage location.  
15

16           25. The computer program product of claim 14 wherein the input set of data  
17 represents auditory speech data from an unknown number of speakers, and further  
18 comprising determining a correct number of speakers from the probability density  
19 modeling the input set of data.  
20

21           26. The computer program product of claim 14 wherein the input set of data  
22 represents image segmentation data from images having regions of different  
23 characteristics.  
24  
25

1           27. A system comprising:

2           a modeling parameter selector selecting a modeling parameter from a  
3 plurality of modeling parameters characterizing a mixture of Student distribution  
4 components;

5           a tractable approximation module computing a tractable approximation of a  
6 posterior distribution for the selected modeling parameter based on an input set of  
7 data and a current estimate of a posterior distribution of at least one unselected  
8 modeling parameter in the plurality of modeling parameters;

9           a lower bound optimizer module computing a lower bound of a log  
10 marginal likelihood as a function of current estimates of the posterior distributions  
11 of the modeling parameters, the current estimates of the posterior distributions of  
12 the modeling parameters including the computed tractable approximation of the  
13 posterior distribution of the selected modeling parameter; and

14           a data model generator generating a probability density modeling the input  
15 set of data, the probability density including the mixture of Student distribution  
16 components, the mixture of Student distribution components being characterized  
17 by the current estimates of the posterior distributions of the modeling parameters,  
18 if the lower bound is satisfactorily optimized.

19  
20           28. The system of claim 27 wherein the lower bound optimizer computes  
21 the lower bound of the log marginal likelihood as a function of prior distributions  
22 of the modeling parameters.

1           29. The system of claim 27 wherein the tractable approximation module  
2 computes a variational approximation of the posterior distribution of the selected  
3 modeling parameter.

4           30. The system of claim 27 wherein one of the plurality of modeling  
5 parameters represents a mean of each of the Student distribution components.  
6

7           31. The system of claim 27 wherein one of the plurality of modeling  
8 parameters represents a precision matrix of the Student distribution components.

9           32. The system of claim 27 wherein one of the plurality of modeling  
10 parameters represents a labeling parameter of the Student distribution components.  
11

12           33. The system of claim 27 wherein one of the plurality of modeling  
13 parameters represents a scaling parameter of a precision matrix of the Student  
14 distribution components.

15           34. The system of claim 27 wherein one of the plurality of modeling  
16 parameters represents a mixing coefficients parameter of the Student distribution  
17 components.  
18

19           35. The system of claim 27 wherein the data model generator generates the  
20 probability density including the mixture of Student distribution components, the  
21 mixture of Student distribution components being characterized by the current  
22 estimates of the posterior distributions of the modeling parameters and an estimate  
23 of the degrees of freedom of each Student distribution component.  
24  
25



1           36. The system of claim 27 further comprising:  
2           a memory storing the current estimates of the posterior distributions of the  
3 modeling parameters.

4           37. The system of claim 27 wherein the input set of data represents auditory  
5 speech data from an unknown number of speakers, and further comprising  
6 determining a correct number of speakers from the probability density modeling  
7 the input set of data.

8           38. The system of claim 27 wherein the input set of data represents image  
9 segmentation data from images having regions of different characteristics.  
10

1           39. A method comprising:  
2           computing a tractable approximation of a posterior distribution for a  
3           selected modeling parameter of a plurality of modeling parameters characterizing  
4           a mixture of Student distribution components based on an input set of data and a  
5           current estimate of a posterior distribution of at least one unselected modeling  
6           parameter in the plurality of modeling parameters;

7           determining whether current estimates of the posterior distributions of the  
8           modeling parameters are satisfactorily optimized, the current estimates of the  
9           posterior distributions of the modeling parameters including the computed  
10          tractable approximation of the posterior distribution of the selected modeling  
11          parameter; and

12          modeling the input set of data by the mixture of Student distribution  
13          components, the mixture of Student distribution components being characterized  
14          by the current estimates of the posterior distributions of the modeling parameters.  
15

16          40. The method of claim 39 wherein the computing operation and  
17          determining operation comprise a first iteration and further comprising:

18          selecting a different modeling parameter from the plurality of modeling  
19          parameters and repeating in a subsequent iteration the operations of computing a  
20          tractable approximation and computing a lower bound using the newly selected  
21          modeling parameter, if the lower bound is not satisfactorily optimized in the first  
22          iteration.  
23  
24  
25

1           41. The method of claim 39 wherein the operation of determining whether  
2 current estimates of the posterior distributions of the modeling parameters are  
3 satisfactorily optimized comprises:

4           computing a lower bound of the log marginal likelihood as a function of  
5 prior distributions of the modeling parameters and a variational posterior  
6 distribution; and

7           determining whether the lower bound satisfies a predetermined criterion of  
8 the selected modeling parameter.

9           42. The method of claim 39 wherein computing a tractable approximation  
10 of a posterior distribution comprises:

11           computing a variational approximation of the posterior distribution.

12  
13           43. The method of claim 39 wherein one of the plurality of modeling  
14 parameters represents a mean of each of the Student distribution components.

15           44. The method of claim 39 wherein one of the plurality of modeling  
16 parameters represents a precision matrix of the Student distribution components.

17  
18           45. The method of claim 39 wherein one of the plurality of modeling  
19 parameters represents a labeling parameter of the Student distribution components.

20           46. The method of claim 39 wherein one of the plurality of modeling  
21 parameters represents a scaling parameter of a precision matrix of the Student  
22 distribution components.  
23  
24  
25

1           47. The method of claim 39 wherein one of the plurality of modeling  
2 parameters represents a mixing coefficients parameter of the Student distribution  
3 components.

4           48. The method of claim 39 wherein modeling the input data comprises:  
5           generating the probability density including the mixture of Student  
6 distribution components, the mixture of Student distribution components being  
7 characterized by the current estimates of the posterior distributions of the  
8 modeling parameters and an estimate of the degrees of freedom of each Student  
9 distribution component.

10           49. The method of claim 39 further comprising:  
11           storing the current estimates of the posterior distributions of the modeling  
12 parameters in a storage location.  
13

1           50. A computer program product encoding a computer program for  
2           executing on a computer system a computer process, the computer process  
3           comprising:

4           computing a tractable approximation of a posterior distribution for a  
5           selected modeling parameter of a plurality of modeling parameters characterizing  
6           a mixture of Student distribution components based on an input set of data and a  
7           current estimate of a posterior distribution of at least one unselected modeling  
8           parameter in the plurality of modeling parameters;

9           determining whether current estimates of the posterior distributions of the  
10          modeling parameters are satisfactorily optimized, the current estimates of the  
11          posterior distributions of the modeling parameters including the computed  
12          tractable approximation of the posterior distribution of the selected modeling  
13          parameter; and

14          modeling the input set of data by the mixture of Student distribution  
15          components, the mixture of Student distribution components being characterized  
16          by the current estimates of the posterior distributions of the modeling parameters.

17  
18          51. The computer program product of claim 50 wherein the computing  
19          operation and determining operation comprise a first iteration and further  
20          comprising:

21          selecting a different modeling parameter from the plurality of modeling  
22          parameters and repeating in a subsequent iteration the operations of computing a  
23          tractable approximation and computing a lower bound using the newly selected  
24          modeling parameter, if the lower bound is not satisfactorily optimized in the first  
25          iteration.

1           52. The computer program product of claim 50 wherein the operation of  
2 determining whether current estimates of the posterior distributions of the  
3 modeling parameters are satisfactorily optimized comprises:

4           computing a lower bound of the log marginal likelihood as a function of  
5 prior distributions of the modeling parameters and a variational posterior  
6 distribution; and

7           determining whether the lower bound satisfies a predetermined criterion.

8           53. The computer program product of claim 50 wherein computing a  
9 tractable approximation of a posterior distribution comprises:

10           computing a variational approximation of the posterior distribution of the  
11 selected modeling parameter.

12  
13           54. The computer program product of claim 50 wherein one of the plurality  
14 of modeling parameters represents a mean of each of the Student distribution  
15 components.

16           55. The computer program product of claim 50 wherein one of the plurality  
17 of modeling parameters represents a precision matrix of the Student distribution  
18 components.

19  
20           56. The computer program product of claim 50 wherein one of the plurality  
21 of modeling parameters represents a labeling parameter of the Student distribution  
22 components.

1           57. The computer program product of claim 50 wherein one of the plurality  
2 of modeling parameters represents a scaling parameter of a precision matrix of the  
3 Student distribution components.

4           58. The computer program product of claim 50 wherein one of the plurality  
5 of modeling parameters represents a mixing coefficients parameter of the Student  
6 distribution components.

7           59. The computer program product of claim 50 wherein modeling the input  
8 data comprises:  
9

10           generating the probability density including the mixture of Student  
11 distribution components, the mixture of Student distribution components being  
12 characterized by the current estimates of the posterior distributions of the  
13 modeling parameters and an estimate of the degrees of freedom of each Student  
14 distribution component.

15           60. The computer program product of claim 50 wherein the computer  
16 process further comprises:

17           storing the current estimates of the posterior distributions of the modeling  
18 parameters in a storage location.  
19  
20  
21  
22  
23  
24  
25

1           61. A system comprising:

2           a tractable approximation module computing a tractable approximation of a  
3 posterior distribution for a selected modeling parameter of a plurality of modeling  
4 parameters characterizing a mixture of Student distribution components based on  
5 an input set of data and a current estimate of a posterior distribution of at least one  
6 unselected modeling parameter in the plurality of modeling parameters;

7           an optimizer module determining whether current estimates of the posterior  
8 distributions of the modeling parameters are satisfactorily optimized, the current  
9 estimates of the posterior distributions of the modeling parameters including the  
10 computed tractable approximation of the posterior distribution of the selected  
11 modeling parameter; and

12           a data model generator modeling the input set of data by the mixture of  
13 Student distribution components, the mixture of Student distribution components  
14 being characterized by the current estimates of the posterior distributions of the  
15 modeling parameters.

16  
17           62. The system of claim 61 wherein optimizer module computes a lower  
18 bound of the log marginal likelihood as a function of prior distributions of the  
19 modeling parameters and a variational posterior distribution, and determines  
20 whether the lower bound satisfies a predetermined criterion.

21           63. The system of claim 61 wherein the tractable approximation modules  
22 computes a variational approximation of the posterior distribution of the selected  
23 modeling parameter.  
24  
25



1           64. The system of claim 61 wherein one of the plurality of modeling  
2 parameters represents a mean of each of the Student distribution components.

3           65. The system of claim 61 wherein one of the plurality of modeling  
4 parameters represents a precision matrix of the Student distribution components.

5           66. The system of claim 61 wherein one of the plurality of modeling  
6 parameters represents a labeling parameter of the Student distribution components.

7           67. The system of claim 61 wherein one of the plurality of modeling  
8 parameters represents a scaling parameter of a precision matrix of the Student  
9 distribution components.  
10

11           68. The system of claim 61 wherein one of the plurality of modeling  
12 parameters represents a mixing coefficients parameter of the Student distribution  
13 components.  
14

15           69. The system of claim 61 wherein modeling the input data comprises:  
16           generating the probability density including the mixture of Student  
17 distribution components, the mixture of Student distribution components being  
18 characterized by the current estimates of the posterior distributions of the  
19 modeling parameters and an estimate of the degrees of freedom of each Student  
20 distribution component.

21           70. The system of claim 61 further comprising:  
22           a memory storing the current estimates of the posterior distributions of the  
23 modeling parameters.  
24  
25